
Re-testing the Perception of Social Annotations in Web Search

Jennifer Fernquist

Google
1600 Amphitheatre Pkwy
Mountain View, CA 94043 USA
jenf@google.com

Ed H. Chi

Google
1600 Amphitheatre Pkwy
Mountain View, CA 94043 USA
edchi@google.com

Abstract

We evaluated the perception of social annotations designed via guidelines recommended by Muralidharan, Gyongyi, Chi, 2012. The initial study found participants noticed the annotation only 11% of the time with annotations shown below the search result snippet. Our refined study revealed that the proposed design with the annotation above the snippet increased noticeability to 60%. Replication studies are often iterative version of old studies, and this was no exception. The new study refined the protocol for measuring 'notice' events, and modified the tasks to include tasks that are more relevant to recent news articles.

Author Keywords

Annotation; social search; eyetracking; user study.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

Introduction

The abundance of information on the web suggests the importance of creating an environment in which users have the appropriate signals to make decisions about which search results are the most useful to them. As more of the web involves social interactions, they produce a wealth of signals for searching the most

interesting and relevant information. Much research has been done on modifying search ranking based on social signals for web pages [1][2][3][5][6], but how should we present the social signals for web search results? The most recent paper that we have found is the CHI2012 paper on social annotations by Muralidharan et al. [4].

Previous Research

Muralidharan et al. [4] studied the perception of social annotations appearing below search results, as in **Figure 1**. Consistent with prior papers, we use the term “social signals” to refer to any social information that is used to affect ranking, recommendation or presentation to the user. We use the term “social annotations” to refer to the presentation of social signals for an explanation as to why a search or recommendation result is presented. Thus, a social signal only becomes an annotation when it is presented to the user.

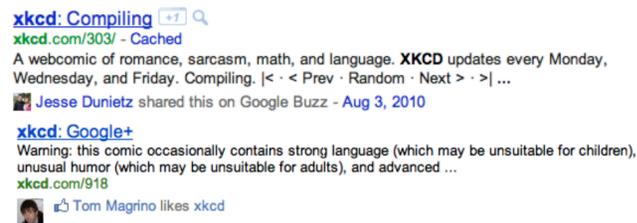


Figure 1: Example of older designs of social annotations. Image is from [4].

Study Protocol

Their first study had two parts: (1) In the first part, participants conducted 18-20 search tasks, randomly ordered. Half were designed so that one or two social annotations would appear in the top four or five results. The search results pages were presented as static mocks

that were generated before the study, customized for each participant.

(2) The second part consisted of a retrospective think-aloud (RTA) where they walked the participant through each task using the eyetrace data post hoc. During the interview, researchers checked noticeability by asking if the participants noticed the social annotations, either by them mentioning they saw them or being explicitly asked if they had seen them. During the RTA the researchers also obtained qualitative feedback about social annotations.

The second study compared the perception of multiple designs of social annotations. They varied profile image size (small, large), snippet length (1, 2, 4 lines), and annotation position (above, below snippet). For this study the same mocks were used for each participant, with customization only for customizing familiar names and faces of people in the annotations. In the second study, noticeability of the annotations was measured by counting the number of fixations.

Findings

In the first study, they found that only 5 of the 45 (11%) of the visible social annotations were noticed. In the second study, they found that there were fewer fixations on annotations when: the snippet length was longer; the image was smaller; and the annotation was below the snippet. They concluded that the optimal design for a social annotation is one with a large picture, above the snippet, with a short snippet length.

Our Method and Replication

We aimed to actually test the proposed annotation design guidelines from study 2 using live user data to

see if people notice the annotations more by using the method from study 1. Specifically, we wanted to test with live data that is relevant to participants (from their connections), as opposed to the static images used previously. An example of a social annotation with the new design is shown in **Figure 2**.



Figure 2: Example result with the new annotation design proposed by prior work. This annotation is above the snippet, has a large image and the snippet is less than 4 lines long.

Study Protocol

Experimental sessions consisted of 3 parts, the first two using essentially the same protocol as experiment 1's in the previous work, with some improvements.

PART 1: SEARCH TASKS

We designed planned 16-20 custom search tasks for each subject, at least eight of which were "social search" tasks designed to organically pull up social annotations. The 8 non-social search tasks were the same as used in the prior work.

In order to ensure that personal results appear for as many queries as required, we designed 2-4 additional social search tasks for each participant that were intended to bring up personal results. This way, if one social search task did not bring up personal results, we gave them the additional tasks to help ensure that they saw 8 tasks with personal results.

PART 2: RETROSPECTIVE THINK-ALOUD

After the search tasks, we immediately conducted a retrospective review of eye-tracking traces for search tasks in which subjects exhibited behaviors of some interest to the experimenter. Review of eye-tracking videos prompted think-aloud question answering about participants' process on the entire task, particular interesting pages, and particular interesting results.

Unlike Experiment 1 in Muralidharan et al. [4], we examined the eyetrace data directly by hand to determine noticeability, rather than through verbal feedback during the RTA tasks.

PART 3: THINK-ALOUD TASKS

Finally, participants performed two or three different search queries for which we determined ahead of time that should bring up relevant personal results. Here we gathered qualitative feedback on social annotations.

Results

In total, we collected eye-trace data for 153 tasks from nine subjects. Each eye-trace data for each task was analyzed by hand by an experimenter to understand: which positions contained personal search results; whether the search result was in the field of view in the browser; and importantly, whether the subject fixated on the result and/or the social annotation. This funnel analysis approach is different than the previous work's approach of asking participants if they noticed the annotations.

We discovered that participants fixated on annotations in 35 of the 58 tasks where they appeared (60%). This is a dramatic improvement over the 11% perception

rate of the Muralidharan et al. [4]. We account this difference primarily to the new annotation design.

Replication Discussion

Access to Previous Experimental Data. We were able to repeat the exact same tasks performed in the previous work but only because we share a co-author who had access to the data. If anyone else tried to replicate the study, they would not have been able to do so as effectively.

Temporal Challenges. Even though the search tasks were identical, because the study was conducted several months later, some of the task questions were no longer topically relevant. For example, one task asked “What is the website for the Google image labeling game?” At the time of our study, the website was no longer active. Similarly, the search task “Find some information about the Nevada law legalizing self-driving cars” brought up news articles from the previous summer, when Muralidharan et al. [4] conducted their research, since it was no longer recent news.

References

- [1] Bao, S., Xue, G., Wu, X., Yu, Y., Fei, B., and Su, Z. Optimizing web search using social annotations. In *Proc. WWW 2007*, 501–510.
- [2] Carmel, D., Zwerdling, N., Guy, I., Ofek-Koifman, S., Har’el, N., Ronen, I., Uziel, E., Yogev, S. and Chernov, S. Personalized social search based on the user’s social network. In *Proc. CIKM 2009*, 1227–1236.
- [3] Heymann, P., Koutrika, G., and Garcia-Molina, H. Can social bookmarking improve web search? In *Proc. WSDM 2008*, 195–206.

This raises a big issue for research replication: changing environments such as time or space. In our case, the tasks lost their relevancy over time. Researchers could help mitigate this by rewriting tasks so they are more relevant but still in the same vein as the original. For example, we could have written a different task that was more topical but would still be categorized as news. It must be decided which would cause the least amount of discrepancy for replication: maintaining the identical, less relevant task or rewriting a relevant task that differs from the original.

Iteration and Refinement. The primary difference in our protocol, measuring perception with fixation data rather than verbal confirmation, offered an improvement to the previous work.

Even with those challenges, we feel that we were successful in our replication efforts. We conducted an almost identical study to confirm the proposed improved design for social annotations and found a large increase in perception.

- [4] Muralidharan, A., Gyongyi, Z., and Chi, E. H. Social annotations in web search. In *Proc. CHI 2012*, 1085–1094.
- [5] Yanbe, Y., Jatowt, A., Nakamura, S., and Tanaka, K. Can social bookmarking enhance search in the web? In *Proc. JCDL 2007*, 107–116.
- [6] Zanardi, V., and Capra, L. Social ranking: uncovering relevant content using tag-based recommender systems. In *Proc. RecSys 2008*, 51–58.